www.bsc.es

Barcelona, 11 January 2017

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Quality Assurance for Multi-model Seasonal Forecast Products (QA4Seas)

Francisco J. Doblas-Reyes
BSC Earth Sciences Department

Objective

QA4Seas aims at **developing a strategy for the evaluation and quality control (EQC) of the multi-model seasonal forecasts provided by the Copernicus Climate Change Service (C3S) to respond to the needs identified among a wide range of stakeholders**.

To achieve the objective the consortium will:

- Consider the evaluation of multi-faceted quality aspects
- Be user driven with a two-stage survey (coordinated with other lots)
- Formulate requirements to the CDS to address user requirements
- Perform a gap analysis of the current information available to users
- Develop a framework and a prototype of the EQC system

Challenges of the QA

- **Addressing adaptation**: it must provide information for climate services and short-term adaptation together because different users approach climate information, where many of them are already familiar with the climate-change problem.
- **Providing consistency**: it must provide information on which trust can be built, requiring a high degree of coherence across products, underlying data sets, processing methods, uncertainty communication, training and guidance, etc.
- **Providing innovation**: it should make operational recent developments from research, with innovative knowledge, methods and technologies to answer real-world issues.
- **Addressing efficiency**: the EQC information should be timely (e.g. available to respond to users' queries with a delay as short as possible), which imposes certain conditions on the approaches and algorithms considered to produce it.

Seasonal forecast QA

- A prediction has no real value without an estimate of its quality based on past performance.
- QA is multifaceted, no single metric is sufficiently comprehensive to characterise the quality of a system or to single-out the best forecast system.
- It addresses administrative (tracking the evolution of systems), scientific (predictability), or socio-economic (users' requirements) questions.
- QA metrics have been developed to study aspects of forecast quality including bias, uncertainty, reliability, resolution, discrimination and sharpness of the forecasts.
- Both deterministic and probabilistic approaches should be considered.
- QA of the European multi-model is not readily available, contrary to the Asian and North American one.

QA and climate services

QA should take place in the context of GFCS, the European Research and Innovation Roadmap for Climate Services, and the Ethical Framework for Climate Services.

Reference points:

- Climate service providers should consider the consequences of their actions for those who may use or be affected by the use of climate service products.
- Climate service products should be open to scrutiny and comparison.

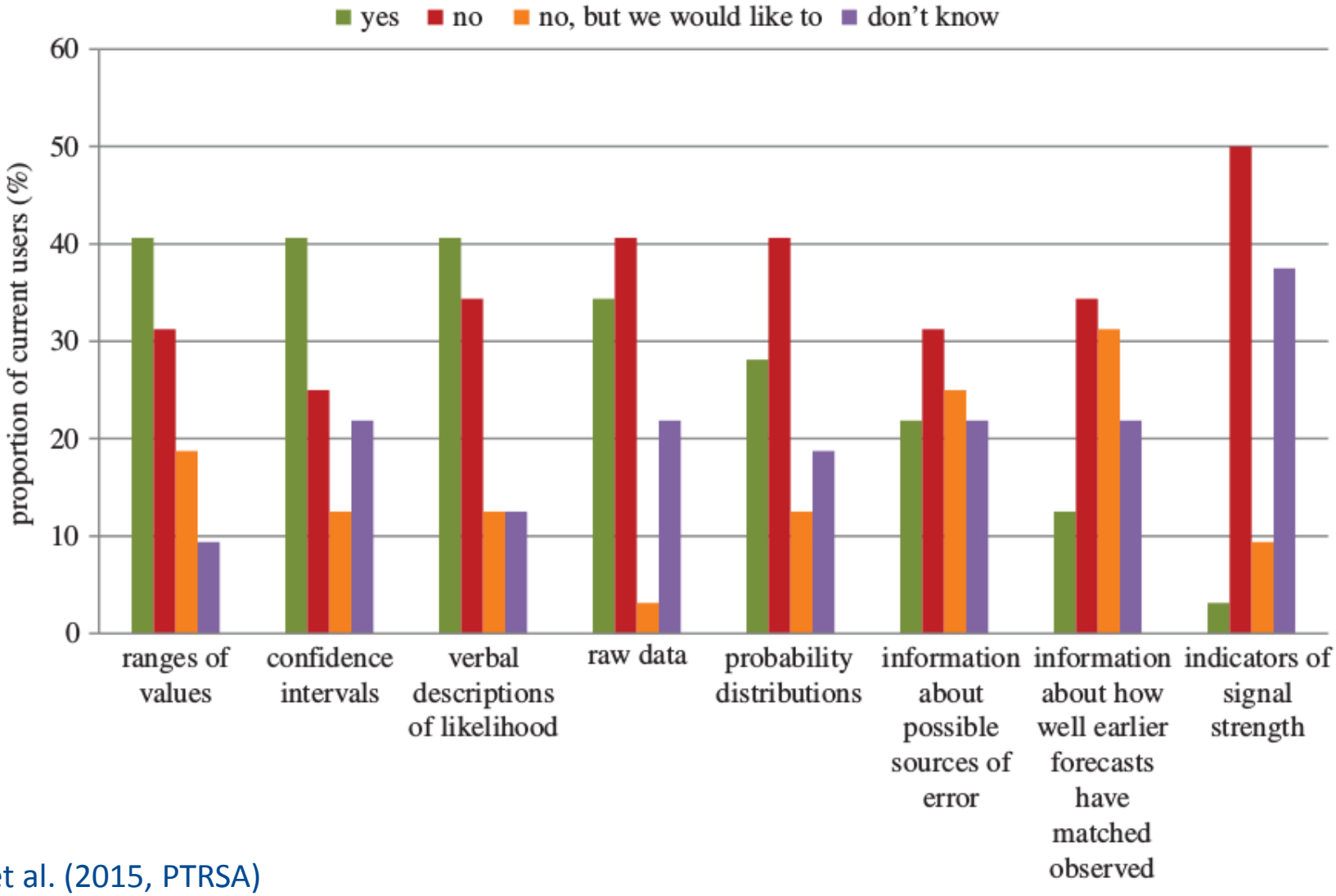
Consortium

Six partners and ten associated participants.

Partner	Nature	Role	Effort (PM)
BSC-CNS	Main contractor	Coordination, data inventory and EQC framework and prototype	88
Univ. Leeds	Subcontractor	Assess user requirements	13.92
Meteoswiss-MCH	Subcontractor	Scientific quality assessment and gap analysis	23
Predictia	Subcontractor	CDS requirements and development of the prototype	28
Univ. Exeter	Subcontractor	Expert statistical advice	3.72
IFCA-CSIC	Subcontractor	Downscaling	14

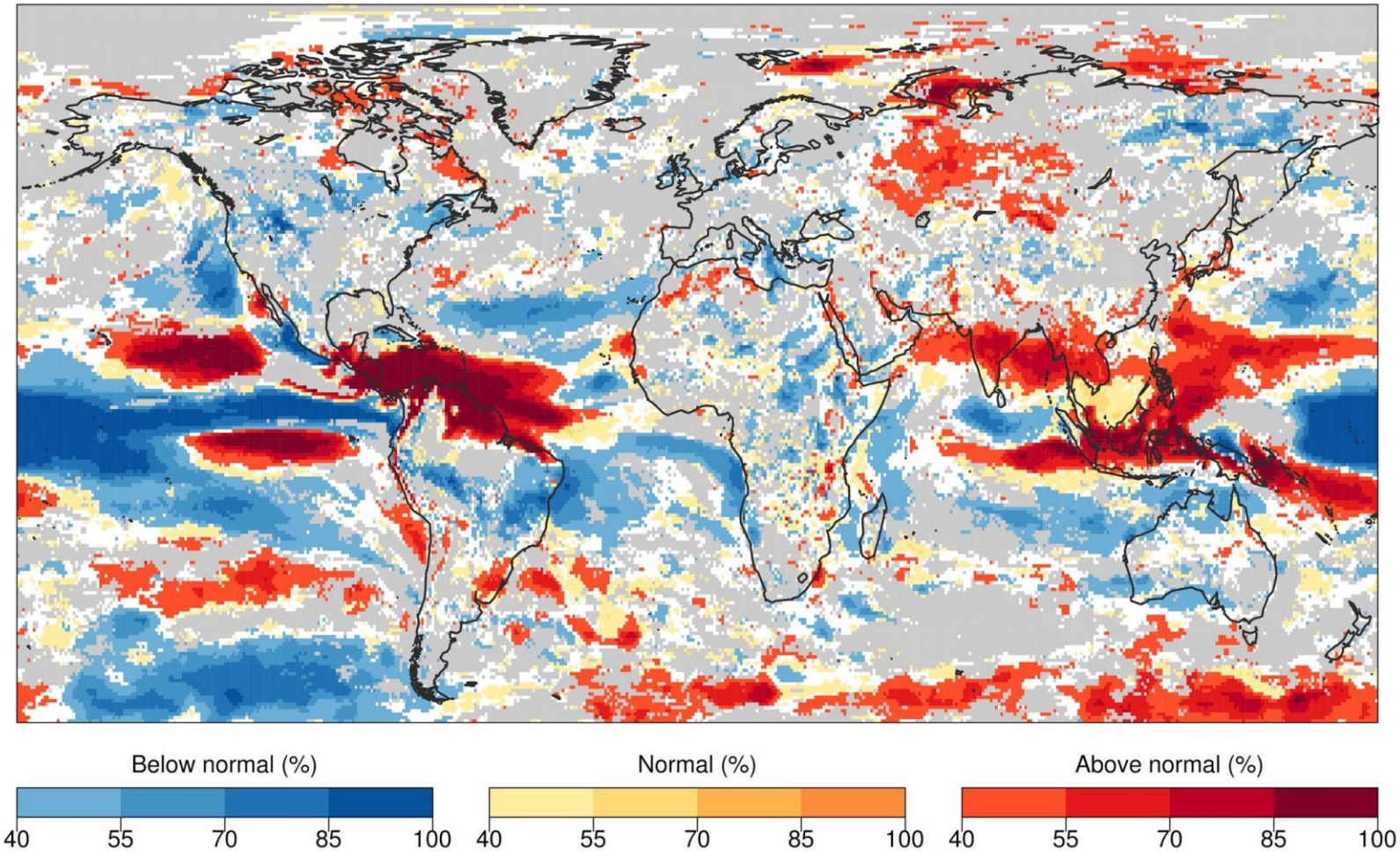
Example of gaps

Proportion of users of seasonal forecasts (n = 32) indicating whether they received different forms of information about uncertainty.



Taylor et al. (2015, PTRSA)

Displaying prediction with verification

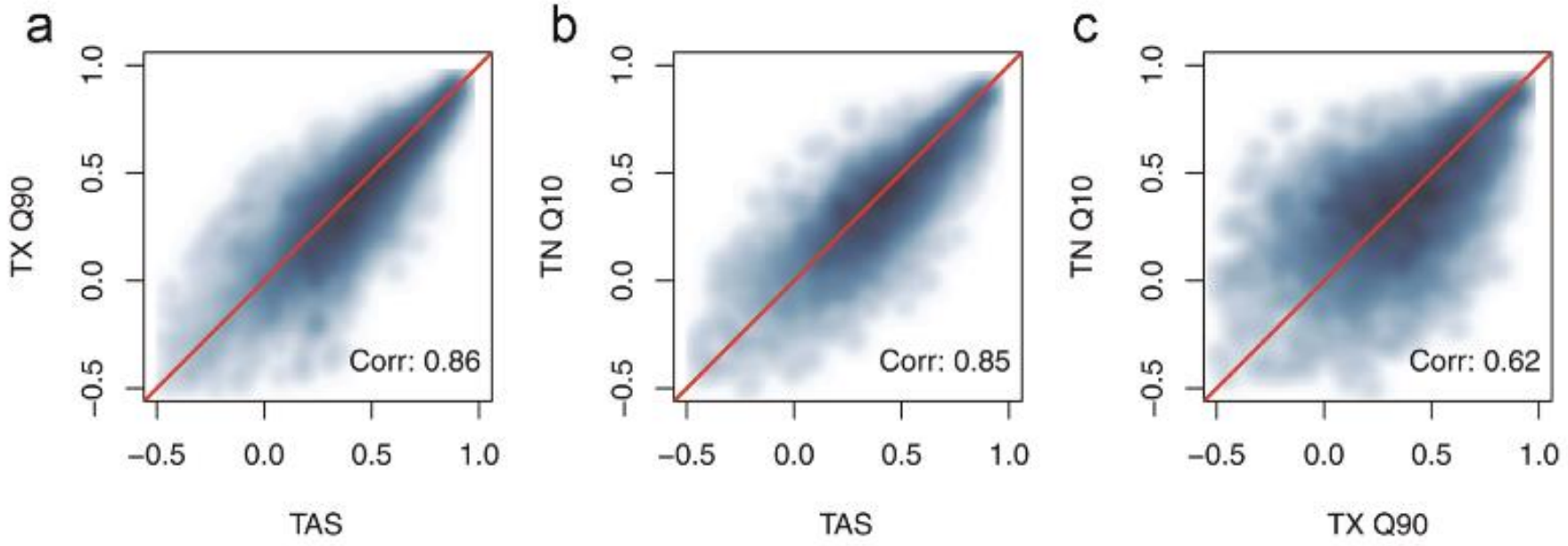


Wind speed prediction for June 1st - August 31st 2015, issued on May 1st 2005.

The most likely wind power category (below normal, normal or above normal), and its percentage probability to occur is shown. "Normal" represents the average of the past. White areas show where the probability is <40% and approximately equal for all three categories. Grey areas show where the climate prediction model does not improve upon the standard and current approach, which projects past climate data into the future.

Verifying “extremes”

Scatter plot of the correlation of the ENSEMBLES multi-model ensemble mean for seasonal-mean daily temperature (TAS) and the 90th and 10th percentiles of Tmax (TXQ90) and Tmin (TNQ90) over the entire globe during JJA over 1979-2005. The correlation between the two samples is in the upper left corner of each panel.

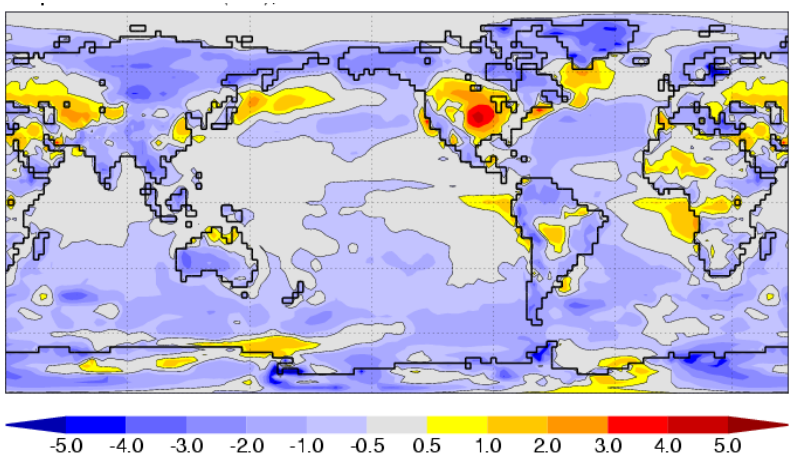


Peppler et al. (2015, WCE)

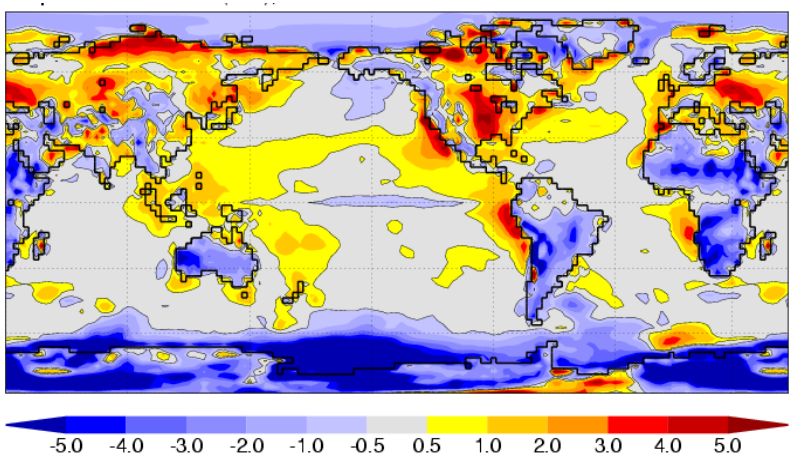
Bias and drift

Mean biases (JJA 2mT over 1993-2005) are often comparable in magnitude to the anomalies which we seek to predict.

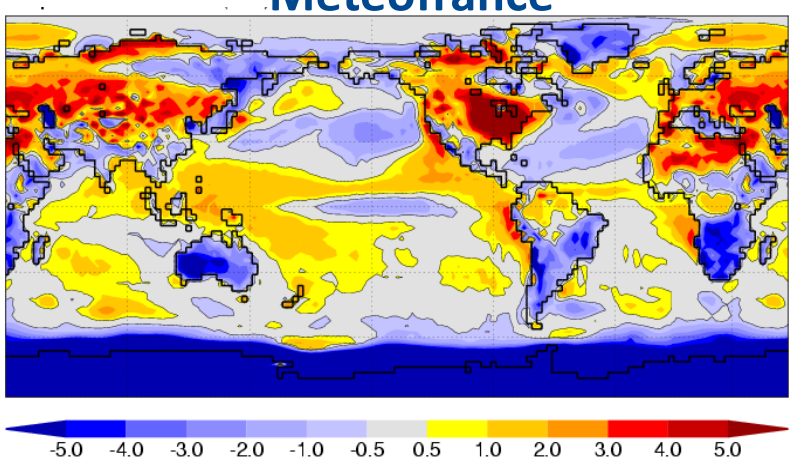
ECMWF



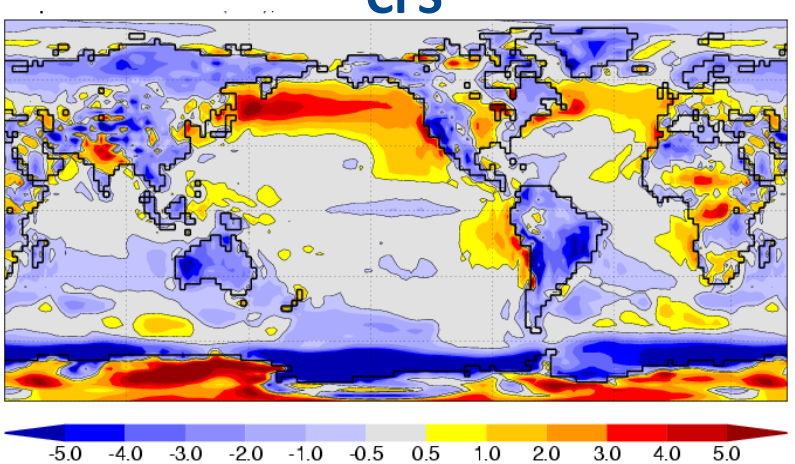
Met Office



Météofrance



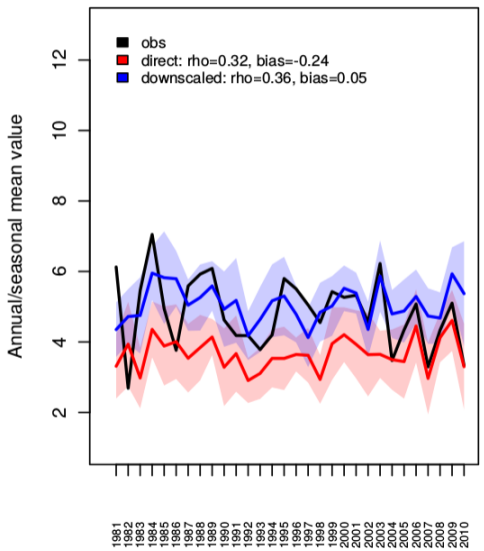
CFS

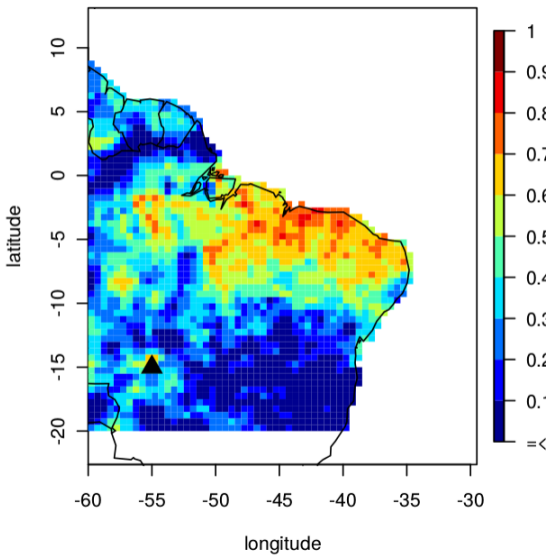


Bias adjustment

- downscaleR is a bias adjustment and perfect-prog statistical downscaling application.
- Performs data manipulation: regridding/interpolation, PCA.
- Bias adjustment with cross-validation: scaling, qq-mapping (various forms), parametric.
- Downscaling with analogs, regression (linear and generalized linear), neural networks, weather typing.
- Parallelization under test.

```
out <- downscale(obs = tp.glb,
                 pred = ncep.pca,
                 sim = tp.system4,
                 method = "glm",
                 cross.val = "loocv",
                 n.pcs = 15,
                 parallel = TRUE,
                 ncores = 3)
# Generate an inter-annual diagnostic plot:
quickDiagnostics(obs = tp.glb,
                 sim = tp.system4,
                 downscaled = out,
                 type = "interannual",
                 loc = c(-55,-15))
```

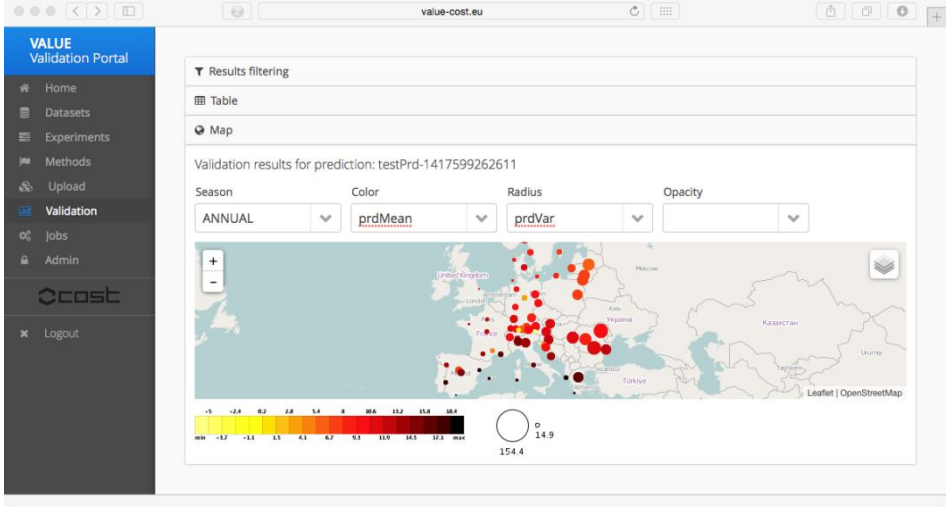




Bias adjustment

- VALUE is an open European network to systematically validate and compare (dynamical and statistical) downscaling methods for climate change research.
- **Metadata** is used in the validation process to decide the suitable validation indices for the different methods. This has been designed in collaboration with ES-DOC and CORDEX-ESD.

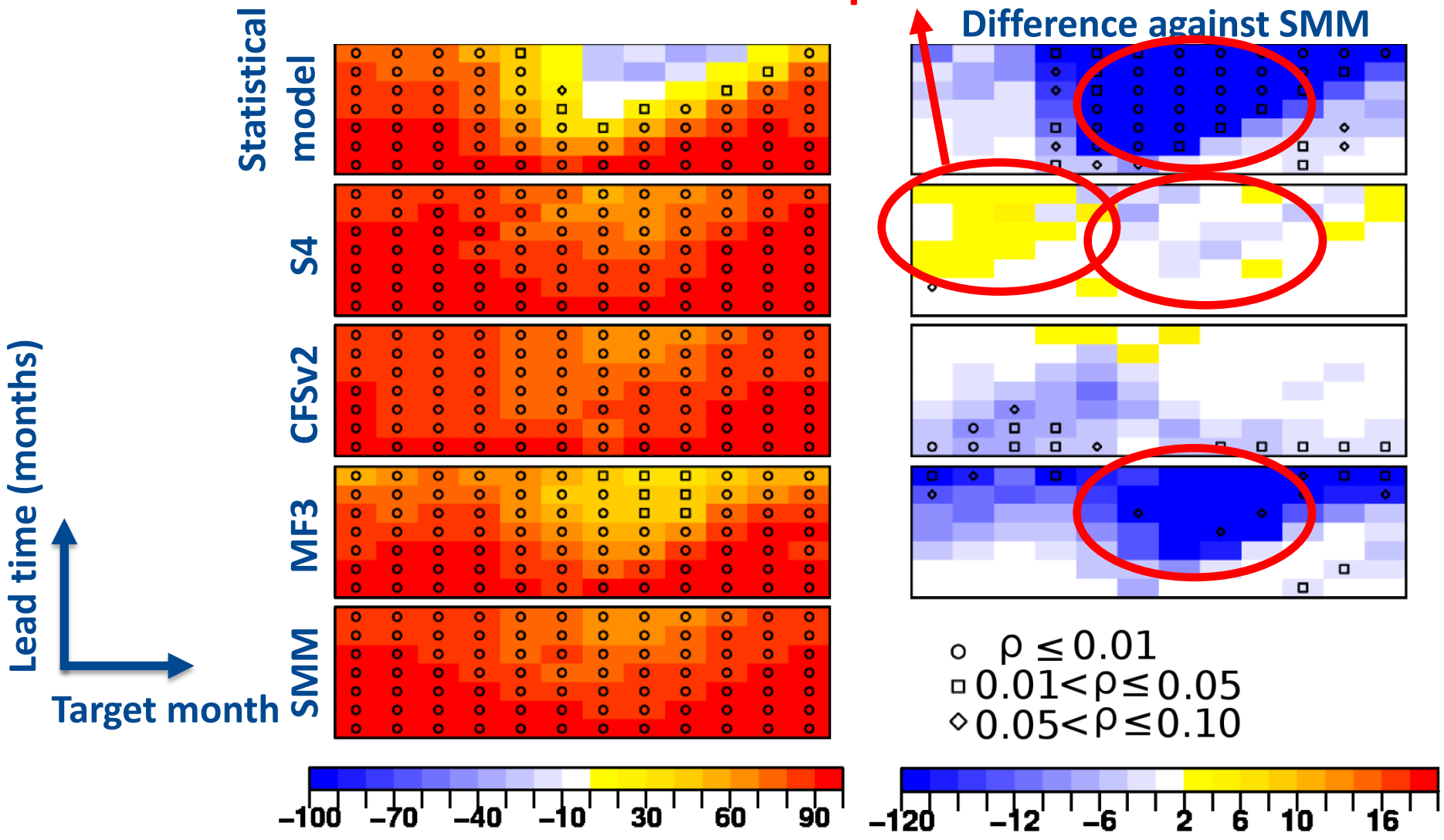
Downscaling **metadata** project,
CIM Control Vocabulary (CV) and schemas.
https://www.earthsystemcog.org/projects/downscalingmetadata/



Multi-model

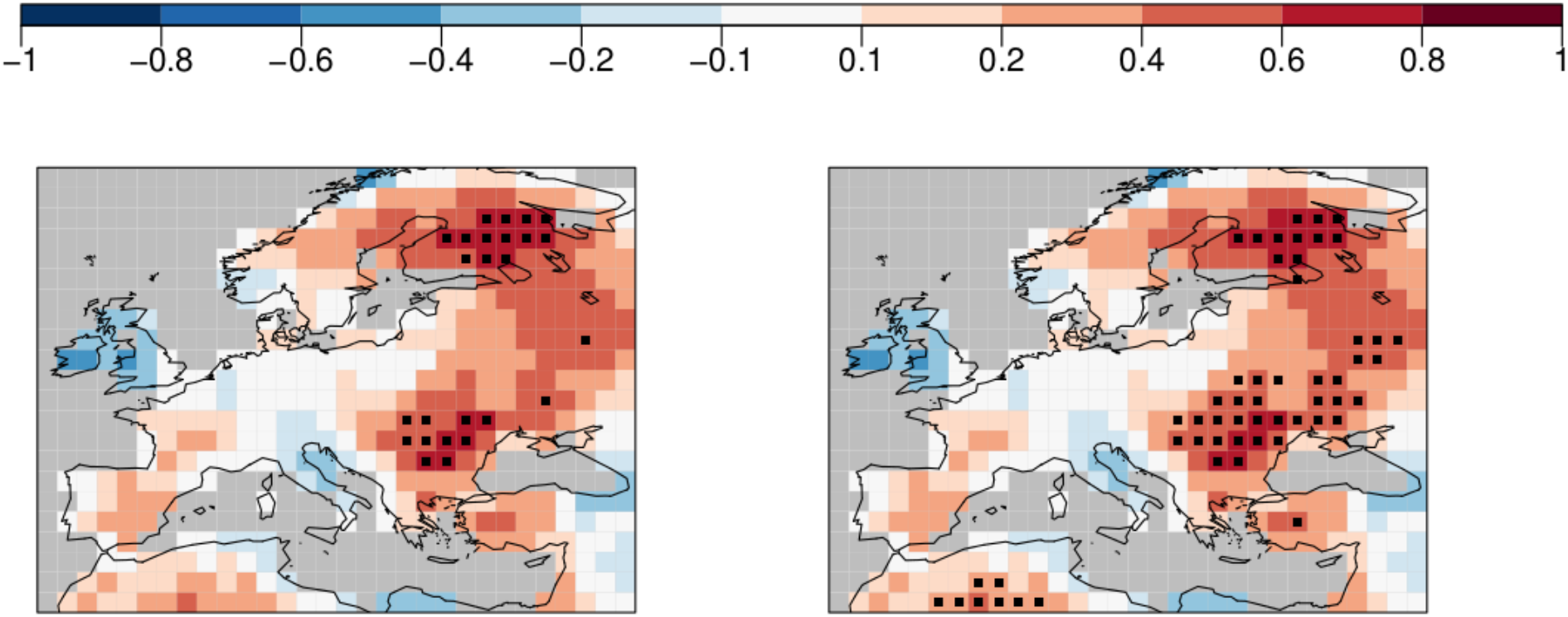
Correlation for the Niño3.4 SST over 1981-2010.

Rodrigues et al. (2014, Clim. Dyn.)



Uncertainty in the QA estimates

Difference in correlation of EC-Earth3 (T511/ORCA025) seasonal one-month lead predictions started every May over 1993-2009 with ERAInt and GLORYS2v1 ics, and internal sea-ice reconstruction with realistic and climatological land-surface initial conditions: left for traditional Fisher test and right for Steiger test (with increased power).

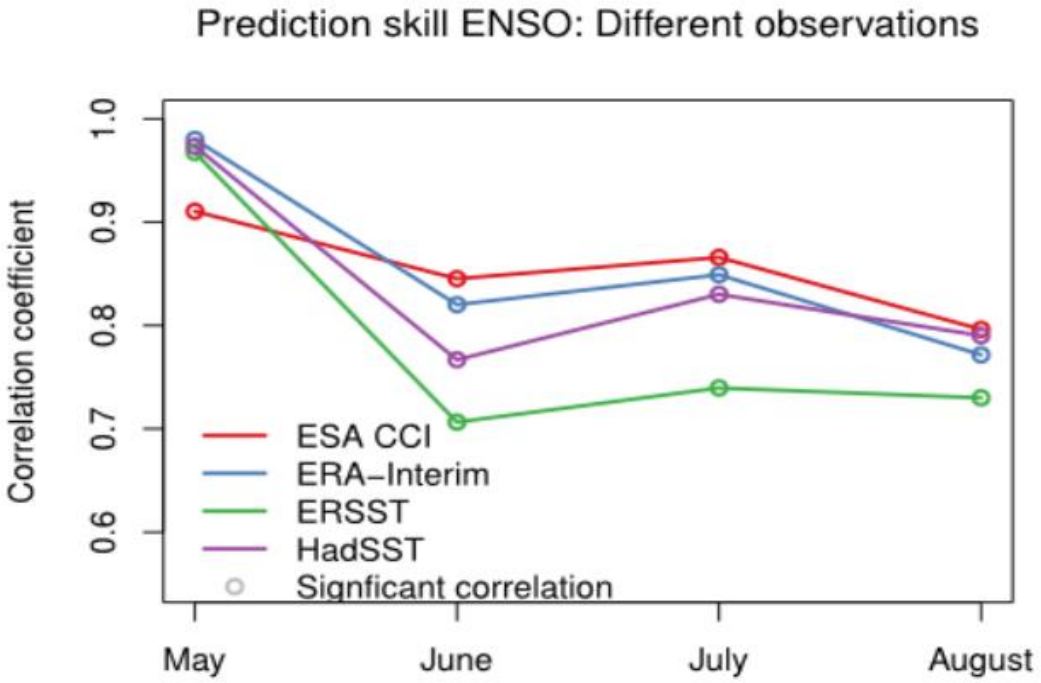


Siegert et al. (2016, MWR)

Observational uncertainty

Need to take into account the large observational uncertainty both in the initialisation and the forecast quality. The Global Climate Observing System (GCOS) is a key actor in this task.

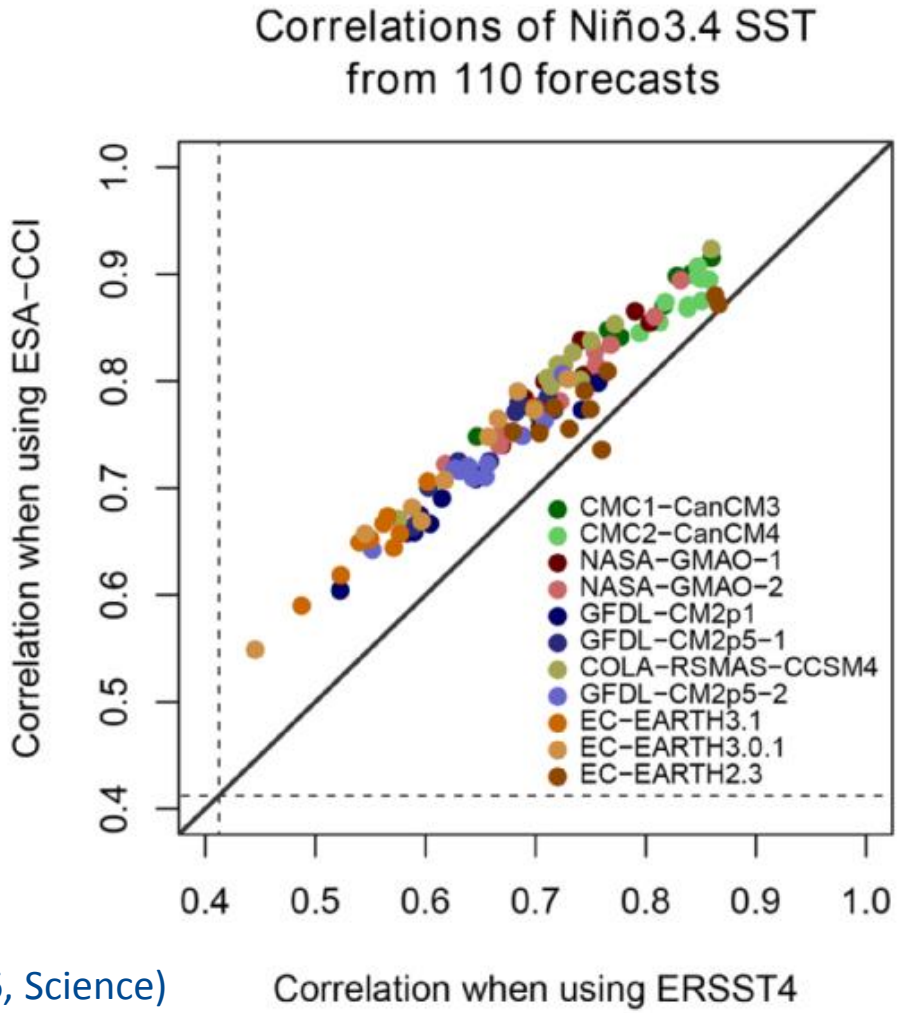
EC-Earth3 (T511/ORCA025) predictions started every May over 1993-2009 with ERAInt and GLORYS2v1 ics, and internal sea-ice reconstruction.



Bellprat et al. (2016, JRSE)

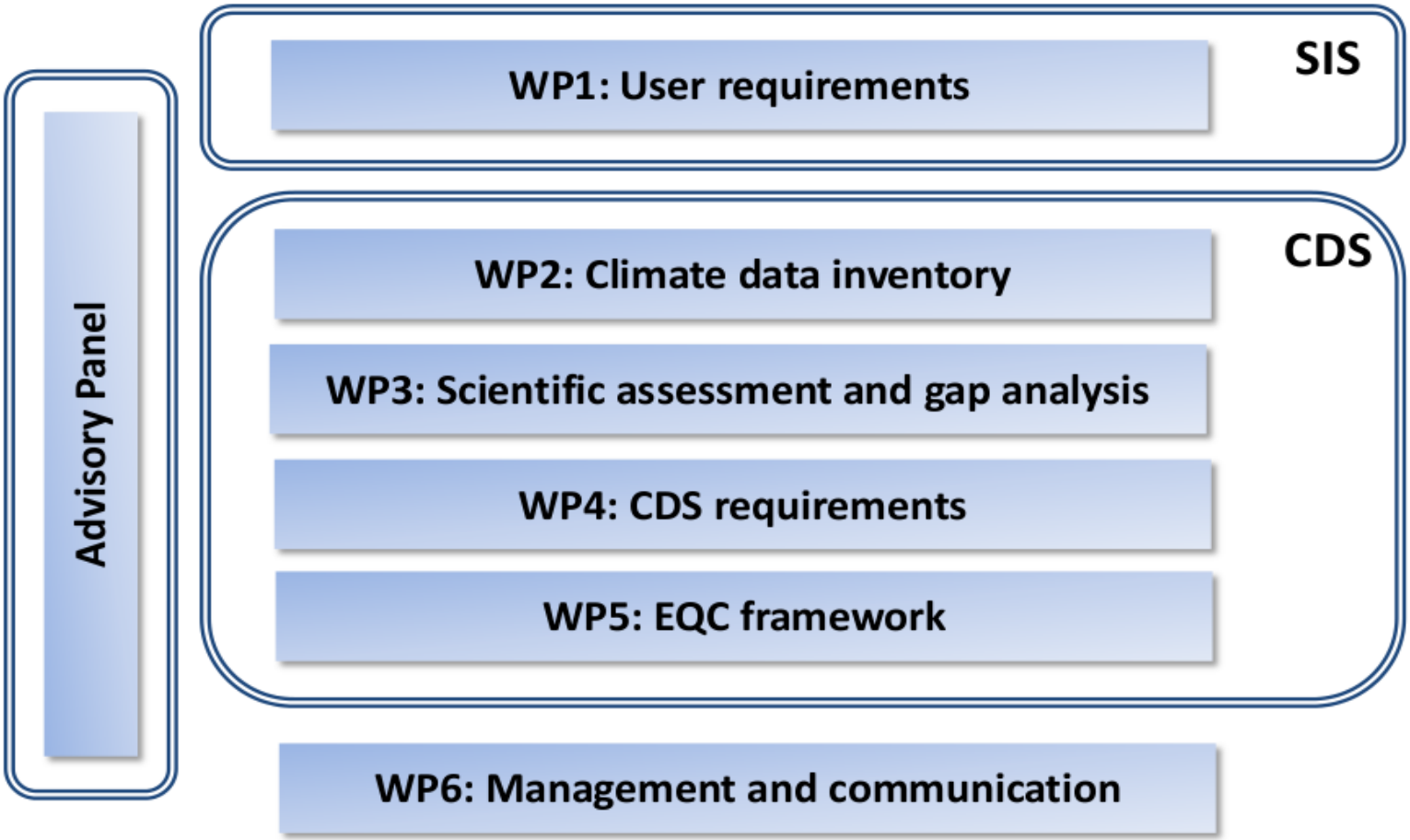
Observational uncertainty

Need to take into account the large observational uncertainty both in the initialisation and the forecast quality. The Global Climate Observing System (GCOS) is a key actor in this task.



Massonnet et al. (2016, Science)

Management


Advisory Panel
WP1: User requirements
SIS
WP2: Climate data inventory
WP3: Scientific assessment and gap analysis
WP4: CDS requirements
WP5: EQC framework
CDS
WP6: Management and communication

WP1: Assessment of users' requirements

- Assess user requirements for the **evaluation and quality control of data** describing the physical climate system in the context of climate services **provision from climate predictions**, and in connection with other applications involving climate information development relevant to C3S.
- Methods used: data collection, surveys, and interviews.
- Workflow:
  - **Strategic framework:** collaboration and coordination with other C3S 51 lots and proof-of-concept projects regarding the users
  - **Collating data** on user requirements from other projects
  - **Survey** for assessing requirements of users involved in the project
  - **Interviews** for assessing requirements of specific users (JRC, EEA)
  - **Survey** for assessing needs of SIS user group and advisory panel
  - **Sharing findings** with other WPs e.g. WP3 scientific assessment and gap analysis and WP5 and EQC framework

WP2: Climate data inventory

- Develop **an inventory of existing climate data sets** to be considered for inclusion in the CDS to address the user requirements identified in WP1. The data requirements of a set of **user-relevant indices** will also be considered to ensure that the necessary ECVs are available.
- Tasks:

1. List of available ECVs to be used in the WP1 survey process (**BSC**, Univ. Leeds).

2. Inventory of additional ECVs and indices from the set of user requirements (**BSC**, Univ. Leeds, IFCA-CSIC, MCH).

3. Coordination with other lots of C3S 51 and with the seasonal prediction data providers of C3S 433 (**BSC**).

WP3: Scientific assessments and gap analysis

- Provide scientific quality assessments for the European multi-model seasonal forecasts serving as an overview of the quality of the pre-operational and POC systems available in the CDS, including an uncertainty assessment of the quality estimates. Carry out a gap analysis of the current quality information available to users, in graphical, text and numerical form, and refer it to the user requirement analysis performed in WP1. Produce recommendations to inform the European research agenda.
- Tasks:

1. Evaluation of forecast quality of pre-operational and POC multi-model seasonal forecasts (**MCH**, BSC-CNS, Predictia, IFCA-CSIC).

2. Assessment of limitations and value added by downscaling of seasonal forecasts (**IFCA-CSIC**, MCH).

3. Gap analysis of information available to users (**MCH**, BSC-CNS).

WP4: CDS requirements

- Produce recommendations for further CDS development based on the user requirements (WP1), the data availability (WP2), and the gap analysis and technical and scientific assessments of the current data sets (WP3) including additional data sets for the CDS, strategies for products and communicate uncertainty, suggestions of user guidance and visualisation requirements, and aspects linked to an efficient data access like metadata standards for the quality information, code development and numerical performance.
- Tasks:

1. Strategies for communicating uncertainty (**Predictia**, Univ. Leeds).

2. Specification of EQC products and standards to be included in the CDS (**Predictia**, BSC-CNS, MCH).

3. Recommendations for computational efficiency and code development guidance (**Predictia**, Univ. Exeter, BSC-CNS, MCH).

WP5: EQC framework

- Produce recommendations for the development of the EQC function for multi-model seasonal forecasts and develop a framework for quality assessment, including a list of key performance indicators for the CDS illustrated with a functional prototype.
- Tasks:

1. EQC framework (**BSC-CNS**, Univ. Exeter, MCH).
2. EQC prototype (**BSC-CNS**, MCH, Predictia, Univ. Leeds).

WP6: Management and communication

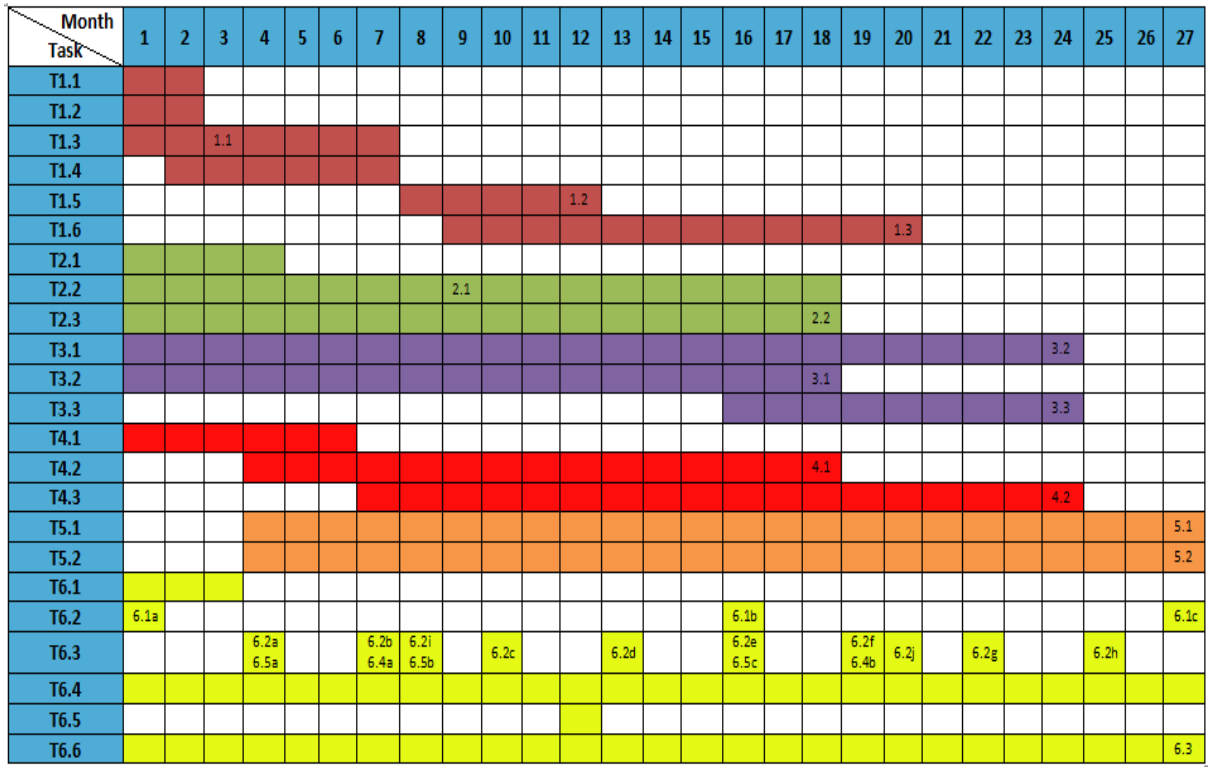
- Ensure the a) management of the project, including reporting to ECMWF, links to ECMWF and other relevant agencies (EEA, JRC), the partners and other relevant C3S contractors and b) the dissemination of the main public outcomes, mainly among the SIS user group.
- Tasks:

1. Description of the service (**BSC-CNS**).
2. Project meetings (**BSC-CNS**).
3. Periodic (quarterly and annual) and final reports (**BSC-CNS**).
4. Link to C3S and coordination with other C3S activities (**BSC-CNS**, Univ. Leeds).
5. Workshop (**BSC-CNS,** Univ. Leeds).
6. Web site and dissemination (**BSC-CNS**).

Project development

The project will be developed in two main phases:

- Service contract 1 (M1-18): User consultation, assessment of user requirements, formulation of requirements and preliminary scientific assessment
- Service contract 2 (M19-27): Gap assessment, and EQC framework and prototype



Existing packages

Developments from SPECS and EUPORIAS. Tried to solve some of the problems identified in the generic "verification" R package.

Vignettes and documentation exist for all of them.

UNIVERSITY OF EXETER

SpecsVerification
- Probabilistic and deterministic scores
- Works on [time x members] arrays

MeteoSwiss

easyVerification
- Applies SpecsVerification scores to arrays of any dimensions, multi-core
- Probabilistic and deterministic scores

SCORES

downscaleR + loadeR
- Data retrieval and homogenization
- Bias adjustment, modes, downscaling
- Probabilistic and deterministic scores
- Visualisation of data and results

s2dverification
- Data retrieval and homogenization
- Bias adjustment, filtering, modes
- Probabilistic and deterministic scores
- Visualisation of data and results

FRAMEWORKS

UC UNIVERSIDAD DE CANTABRIA CSIC

Existing packages

These valuable tools should be considered as the building blocks:

- Select best features of the packages
- Adapt all scores to work on [time x members] arrays and request work to increase performance in the computation of scores
- Define a generic data structure (any tool could be applied to all data)
- Fundamental performance issues:
  - RPSS of a 2 GB array (global map) on 1 core takes ~2 hours (on a platform P)
  - easyVerification can run on N cores, speed-up=~ N/2 (limited by data transfer), and takes ~30 min. on 8 cores (on same platform P)
  - S2dverification splits the array in chunks of grid points and forecast horizons, gathering results in an 'ff' file, with speed-up =~N (data transfer avoided), and takes ~15 min. on 8 cores (on the same platform P)

Links to other contracts

ITT	Information required in this proposal	Information provided by this proposal
Multi-model seasonal forecast providers (C3S 433)	Technical characteristics (timeliness, specific formats, initialisation methodology, etc.) of the seasonal forecast systems contributing to the CDS	Forecast quality assessment, including drift and systematic error characteristics, of each individual forecast system, as well as an assessment of the strengths of the multi-model approach
Software Infrastructure for the CDS (C3S 23)	Methodology to access the CDS data. Overview of all ECVs available. Data availability and quality estimates. Metadata standards. Methodology for integrating relevant standards and tools into the CDS.	Timeliness required of relevant ECVs (observations, reanalyses and single and multi-model seasonal forecasts).
Software Development of the CDS Toolbox (C3S 25)	Tools used for sub-setting, re-gridding and format conversion. Data model to be used to combine data and products of different types. List of functions offered by the toolbox. Feedback from the “Big Data” techniques investigated. Methodology for integrating relevant standards and tools into the CDS.	QA4Seas quality assessment prototype. Quality assessment dissemination and visualisation standards. Recommendations for computational efficiency of the quality assessment.
EQC for SIS (C3S 441, 52)	Communicate user requirements, gaps and market analysis; list of stakeholders with interest in seasonal prediction. Proof-of-concept and pre-operational EQC assessment reports for all C3S sectors. List of sectoral climate information indices.	Relevant metrics, key performance indicators and illustrative examples for the use of user relevant ECVs and indices from multi-model seasonal forecasts. Examples of EQC metrics for sectoral climate impact indicators for the applications that will be included in the CDS. Methodology to estimate and validate the climate information indices from the CDS.